R E P O R T R E S U M E S

ED 003 600 56

THE RELATIONSHIPS OF INTEREST AND ATTENTION TO RETENTION AND ATTITUDE CHANGE.

BY- BECKER, SAMUEL L.
IOWA UNIV., IOWA CITY

REPORT NUMBER NDEA-VIIA-739 PUB DATE 63
EDRS PRICE MF-$0.09 HC-$1.88 47P.

DESCRIPTORS- *GRADE 5, *TELEVISED INSTRUCTION, STUDENT REACTION, *STUDENT ATTITUDES, STUDENT BEHAVIOR, FILMS, *CHANGING ATTITUDES, *ATTENTION SPAN, *RETENTION STUDIES, IOWA CITY, IOWA, GALVANIC SKIN RESPONSES (GSR)

THE FOLLOWING QUESTIONS WERE EXPLORED IN THIS STUDY--(1) WHAT IS THE RELATIONSHIP AMONG RESPONSES OBTAINED WHEN SUBJECTS VIEW AN INSTRUCTIONAL TELEVISION PROGRAM AND NOTE THEIR INTEREST BY PRESSING OR RELEASING A BUTTON, BY CHECKING COLUMNS ON A SHEET OF PAPER, AND WHEN GALVANIC SKIN RESPONSES (GSR) ARE CHARTED, (2) WHAT IS THE RELATIONSHIP BETWEEN THE AMOUNT OF INFORMATION RETAINED FROM THE PROGRAM AND TYPE OF RESPONSE PROFILE, (3) CAN GSR BE USEFUL TO PREDICT THE AMOUNT AND DIRECTION OF ATTITUDE CHANGE FROM THE PROGRAM. A MINIMUM OF 30 FIFTH-GRADE STUDENTS WERE ASSIGNED TO EACH OF 3 EXPERIMENTAL GROUPS. STIMULI WERE TWO HALF-HOUR TELEVISION PROGRAMS. EACH SUBJECT WAS ALONE IN A SMALL VIEWING ROOM AS HE VIEWED EACH PROGRAM. HE WAS TESTED ON THE FACTUAL CONTENT OF THE PROGRAMS AND ATTITUDES TOWARD CONTENT IMMEDIATELY BEFORE AND AFTER EXPOSURE. THE PAPER AND PENCIL AND PUSH-BUTTON METHODS OF AUDIENCE RESPONSE TO FILMS RESULTED IN SIMILAR PROFILES. WHEN THESE ARE COMPARED TO THE GSR METHOD, HOWEVER, THIS WAS NOT THE CASE. GSR'S TENDED TO BE EVOKED BY THOSE PARTS OF THE PROGRAM WHICH WERE THOUGHT "BORING." NONE OF THE METHODS WERE GOOD PREDICTORS OF RETENTION. GSR RESEARCH IN PERSUASIVE COMMUNICATIONS IS RECOMMENDED IN LIGHT OF RELATIONSHIPS OF ATTITUDES TO GSR SCORES. (JC)

ED003600

#739
739

THE RELATIONSHIPS OF INTEREST AND ATTENTION TO RETENTION AND ATTITUDE CHANGE

SAMUEL L. BECKER, Director
Division of Television, Radio, Film
University of Iowa

The University of Iowa
Iowa City, Iowa

1963

U. S. DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
Office of Education
This document has been reproduced exactly as received from the person or organization originating it. Points of view or opinions stated do not necessarily represent official Office of Education position or policy.

Addenda: "The Relationships of Interest and Attention to Retention and Attitude Change"

There is one point made in the attached report which needs modification, another for which additional information will help interpretation.

On pages 17-18 and 34-35, we have indicated as untenable the conclusion that GSRs are evoked more by boring than by interesting content. Though we would still hold with this conclusion in the context of this study, it should be noted that evidence has been found in other studies that excessive boredom, such as that occurring in sensory deprivation experiments, can lead to an increase in GSRs. An individual in a situation in which he cannot sleep, and yet in which he receives no stimulation, will tend to become restless and agitated. As Berlyne has noted (D. E. Berlyne, Conflict, Arousal, and Curiosity. New York: McGraw-Hill, 1960. P. 190.), "We may . . . surmise that sensory deprivation becomes aversive when internal factors cause a rise in arousal and the lack of stimulation renders the cortex incapable of keeping arousal within bounds . . . . Internal stimuli, e.g., from thoughts, may . . . contribute to endogenous arousal, since they are so prone to emerge in the absence of external foci of attention in normal, waking human beings."

On pages 13-16, we reported on the reliability of the various measures of audience response. However, since various means were used for estimating reliability of the three audience response methods, it was not possible to make a very satisfactory comparison of their reliabilities. Since this report was printed, we have recalculated the reliabilities, using a technique which permits more meaningful comparisons. This is the intraclass correlation method (R. L. Ebel, "Estimation of the Reliability of Ratings," Psychometrika, (1951), 407-424). By this method, one can obtain an estimate of average reliability for a given number of raters and an estimate of the reliability of a single rater. In the present instance, each subject in a group was treated as a rater who rated each shot of a program as interesting or not interesting. The obtained reliabilities are cited below, along with the number of cases upon which each average reliability coefficient was based.

Reliabilities of each Method of Response by Groups and Programs

	Constitution Program			UN Program		
	Ind. reliab.	Ave. reliab.	n	Ind. reliab.	Ave. reliab.	n
GSR - Group I	.16	.68	11	.14	.81	26
Group II	.20	.86	25	.06	.55	21
Group III	.16	.80	21	.18	.82	22
Push-button	.16	.80	22	.12	.78	25
Paper and Pencil	.06	.64	26	.03	.44	30

It is readily apparent that the GSR and push-button method are of similar reliability. The paper and pencil method, on the other hand, is less reliable.

THE RELATIONSHIPS OF INTEREST AND ATTENTION TO RETENTION AND ATTITUDE CHANGE

Samuel L. Becker, Director
Division of Television, Radio, Film
University of Iowa

The research reported herein was supported in part by a grant from the U. S. Department of Health, Education, and Welfare
Office of Education

The University of Iowa
Iowa City, Iowa
1963

There are many persons whose help in this study must be acknowledged. Foremost is Mr. Gary Cronkhite, whose ideas and assistance helped to overcome and avoid many difficulties. The willing and intelligent work of Mr. David Couch was indispensable. Advice and assistance with problems of physiological measurement came from Prof. Harold Shipton, director of the medical electronics laboratory of the College of Medicine. There were others, too numerous to mention, but including Prof. Lloyd Smith, Mrs. J. Igram, Mr. G. J. Wolfe, Mr. and Mrs. D. Lynch, Mrs. G. Cronkhite, and Mrs. R. Barnes.

Introduction

One of the major disagreements among those planning and producing educational films and television programs today has to do with the relationship between "interest" in a program and retention of information or attitude change resulting from the program. To put it in the terms in which it is usually argued, the question is how important is "showmanship" in an educational program. This study was designed to explore this relationsiip and, more specifically, to examine the relative fruitfulness of various procedures which may be used for research on the problem.

A second and related problem with which this study dealt was one revolving about methodological problems of assessing audience interest in various sections of an educational program. I have discussed the importance of such methodological research elsewhere. The gist of the argument is as follows:

> In the history of science, one of the facts which stands out is the interdependence of growth in knowledge and perfection of tools or methods for investigation. In instance after instance, expansion of knowledge has needed to wait upon new research methods or an improvement in or better understanding of the available methods....
>
> One of the problems which has concerned the producer of various kinds of mass communications for many years has been how to pre-test his message. The production of a series of radio or television programs or films is an expensive matter, not only in the expenditure of money but in time and effort. In many cases, the costs of distribution are even greater than those of production. Because of these factors, there has been strong motivation for finding some means of maxi-

mizing the probability of attaining the desired goals... This meant that some method was needed for studying the reactions to discrete elements within a program, rather than the reactions to their sum. It was not good enough to find out whether people did or did not like a particular television program. The producers needed to know very specifically what it was that was not liked, what was liked, and the reasons in both cases so that weak or mediocre elements could be deleted or made more effective. Merely exposing an audience to a program and then interviewing the audience members seemed to result in only vague or general comments. ... The solution to this problem seemed to be some means of securing from each member of the test audience a continuous reaction to the program as it was in progress. With this "profile" of reactions as a guide, interviewers could probe for the reasons for specific reactions. In some cases, the sections of programs which received notably high or low responses could be replayed to aid the recall of the respondents. In any case, the profile could facilitate focused interviews about reactions.

Many methods of securing these continuous reactions or profiles have been used.... The difficulty with all of.../them/ has been the paucity of research on the methods themselves. Almost every researcher in the country who is carrying out research on reaction profiles is using either a different type of device or a different procedure or, in most cases, both. Little is known of the effect of different types of procedures with different audiences and different subject matter. Little is known of the effect of using different types of devices. In applied research, where one's only goal is the improvement of a specific program or series of programs, this may make little difference /except that such matters as the relationship between "liking" or "interest" and retention or attitude change would seem important to know/. However, as we attempt to make broader generalizations which have meaning beyond a particular series, we need to know whether all procedures used to secure the data upon which generalizations are based are measuring the same or different concepts. The goal of our research in this area at the State University of Iowa is to gain a better understanding of these various methods through experimental comparisons of procedures and devices.[1]

In research on this general subject to date, we have shown the

way in which audience profiles are affected by the degree of effort required for subjects to respond either negatively or positively to a communication,[2] the relationship between push-button and paper and pencil responses,[3] and the relationship between children's interest in a television program and the predictions of this interest by educators and parents.[4] The only other profile research to date which seems to be closely related to the subject of the present study is that by Twyford.[5] Most important, in relation to the present study, was the indication in Twyford's results that a negative relationship may exist between "liking of" and "learning from" a film. If substantiated by other research, this fact has quite important implications for the producer of educational films and television programs. Aspects of this relationship, as I indicated earlier, were explored in the present study.

The third problem addressed by this study was that of using the galvanic skin response (or GSR) as a measure of interest or attention.

The importance of finding a means to measure <u>attention</u> or <u>focus</u> upon various aspects of television programs and films is so great that every possibility should be carefully and exhaustively tested. Before we can do a very good job of experimenting with the learning efficacy of various production elements in these programs, we need to be able to separate two of the major factors which may account for lack of learning. Students may fail to learn

from a particular program unit because their attention was not held to the program and they were attending to something else, or they may fail to learn because of an inadequacy in the program unit. As an example of the latter, the stimulus in the program unit may be too dissimilar to the stimulus in the testing situation to make generalization possible. In order to improve our chances of understanding why learning fails to occur, we need to be able to separate those instances in which it failed because of lack of attention from those in which it failed despite the presence of attention.

In addition, though it further complicates matters, we must ultimately distinguish among two components or aspects of attention, fixation on a stimulus and level of tension. These are not necessarily related. An individual can obviously fixate or focus upon any part of a field and, independently, can be any place on the continuum from a relatively relaxed state to a highly tense state. Optimum learning from a program should be facilitated when attention is focused upon the desired part of the field and when the learner is in a relatively high state of tension (drive).

For readers unfamiliar with the GSR, a brief explanation may be desirable. It has been recognized, since the late 19th century, that an electrical current, or a change of potential, can be detected if a sensitive galvanometer is connected to two points of the skin, and the person so connected is subjected to stimuli

which evoke affective responses. GSR is the decrease in skin resistance measured by an increase in current. It is the change in resistance, not the level of resistance. The most generally accepted explanation for this response is that "nervous impulses responsible for sweat secretion create concomitantly in the sweat gland membranes a condition of increased permeability to current passing through the galvonometer circuit."[6] Researchers have inferred various psychological states from this response. In general, it is taken to be either an indicator of affect or of general energy mobilization. Woodworth says that it is a measure of sympathetic nervous system activity.[7] Freeman says that "changes in electrical resistance provide an extremely delicate index of the 'vigilance' of the central nervous system at any given time."[8] Morgan and Stellar carefully state that "it is hard to say just what the galvanic skin response does and does not indicate, but it certainly is a very sensitive sign of something going on in the nervous system... A good general statement about it is that it comes out whenever a stimulus or idea sets a person in readiness or expectancy of something happening."[9]

There is some evidence that indicates that the GSR should be useful as an index to the portions of television programs or films which are most meaningful to the audience. For example, Woodworth has found that "mechanically minded men gives PGR [what we are calling GSR] at sight of machines in action, a child at sight

of animals playing, an adolescent girl at fine clothes."[10] Bingham found that words which subjects rated higher in terms of their personal meaningfulness, significance, and importance evoked the highest average change in skin resistance.[11] That the utility of GSR for research on "meaningfulness" is not a foregone conclusion is clear when one considers the other behaviors which tend to be accompanied by increased GSRs. As Woodworth and others have indicated, these are energetic movement and the preparation for such movement, deep breathing, coughing, sneezing, sighing, laughing, clearing the throat, and yawning.[12]

Thus, we need to discover whether the GSR responses to our planned stimuli, such as television programs or films, can be identified sufficiently well to make this a useful tool for communication research.

Research Objectives

The following questions were explored in this study:

1. What is the relationship among response profiles obtained when fifth grade students view an instructional television program and note their "interest" reaction continuously by pressing or releasing a button, when they note their reactions each 30 seconds by checking the "interest" or "disinterest" column on a sheet of paper, and when their galvanic skin responses are charted con-

tinuously while the students are viewing a program? For which types of stimuli do these methods provide similar results? For which do they provide different results?

2. What is the relationship between the amount of information which students retain from the programs and each of these types of response profiles? Which method of responding best predicts the information retained?

3. Can GSR be useful to predict the amount and direction of attitude change from the program?

Procedures

Subjects were fifth grade students. A minimum of 30 were assigned to each of three groups. However, because of failure to complete one or more parts of the test battery correctly, a number of subjects were eliminated. All were retained, however, for whom enough data were available to be usable in one or more of the analyses. Thus, analyses were based upon a maximum n of 31 in Group I, 27 in Group II, and 32 in Group III.

Stimuli were two half-hour television programs, both of which were produced as part of a fifth grade social studies television series. One was "The Constitution of the United States of America." Foci of this program were the Articles of Confederation, the Constitutional Convention, the Bill of Rights and many of the men who were involved in their development. The other

program was the "Story of the United Nations." Focus in this program was the work of UNESCO, UNICEF, and WHO. Both programs were essentially illustrated lectures, typical of a large proportion of the educational television programs currently used. The same teacher was used on both programs. Both programs also employed a large number and variety of visual devices.

Each subject was exposed to the program on the Constitution and, approximately one week later, to the program on the United Nations. He was alone in a small film viewing room as he viewed each program. He was tested on the factual content of a program and attitudes toward that content immediately before and immediately after exposure. It was planned initially to do the pre-testing one month before each subject viewed the programs. Because it was necessary to run subjects over such a long period of time, and because it was found that this introduced additional error variance (changes of attitude and knowledge between pre-testing and exposure periods), the pre-test was shifted to the time immediately before exposure.

The design for the measurement conditions is shown in the six-fold table on the following page.

		Program	
		Constitution	United Nations
Group	I	Push-button and GSR	Paper & pencil and GSR
	II	Paper & pencil and GSR	Push-button and GSR
	III	GSR	GSR

A Gilson Medical Electronics Model M8PM Macro-Polygraph was used for recording GSR and push-button responses. This polygraph made it possible to continuously record on a single moving graph the GSRs, heart rate (which is not being reported here), push-button responses, and pre-determined cue marks which permitted the later identification of each segment of a graph with the associated segment of the television program to which the student was responding at that moment. A graph speed of 2.5 centimeters per second was used. The criterion measure for GSRs was simply the presence of GSRs -- changes in resistance -- rather than base level or even the extent of resistance changes.

Silver electrodes were used for obtaining the GSR measures. One was taped to the left palm and one to the back of the left hand of each subject. Electrode jelly was used to insure a better contact between electrode and skin. The subject was asked to keep his left hand resting easily on the left-hand arm rest of his seat and to relax during the showing but to move as little as

possible. A low light level and quiet were maintained in the viewing room for a few minutes prior to the start of the film so that the subject could become accustomed to the situation and relax as much as possible.

As indicated in the table above, GSRs were measured for all subjects during both film viewings.

Push-button respondents were asked to push a toggle switch against a slight spring resistance when they did not find the program interesting and to release it when they became interested. Paper and pencil respondents were asked to check a paper chart each 30 seconds when the appropriate number was flashed beside the screen upon which they were viewing the television program; on the right if they were interested, on the left if they were not. The instructions given for each type of measurement and a copy of the paper and pencil chart may be found in Appendices A and B.

The retention examinations and the attitude instruments were developed in cooperation with Associate Professor Lloyd Smith of the University of Iowa College of Education, who is educational consultant on the social science series from which the stimulus programs were selected, and Mr. Edwin Harbeston, the teacher on the programs. They were based upon transcripts of the programs which included descriptions of both audio and video elements.

One of the serious problems encountered in this study was the number of factors which could prevent the obtaining of a valid GSR profile. Because of this problem, in large part the result of inadequate pre-testing with the particular polygraph used in the study, a large number of GSR graphs had to be discarded.

The response profiles which were retained were tabulated in a number of ways. The first was by retention test items. For this tabulation, we noted the section (or sections) in a program when the information needed for a correct response to an item was being transmitted. The subjects in the paper and pencil group who indicated interest by an appropriate paper and pencil response during this section were then counted. The subjects in the push button group who were not pressing the button during this section were also counted. And the subjects who emitted at least one GSR during the section were counted. For this last, of course, the 1.5 to 2 second latency period which one gets between a stimulus and the resultant GSR was taken into consideration. For all methods, when the information for an item was given more than once, all such sections were treated as one. The result of this tabulation was three profiles for each program, showing the number of subjects indicating interest (if we define the GSR as an indicator of interest) by each of the response methods during each section in which critical information was being transmitted.

There were 25 items (sections thus tabulated) for the United Nations program; 33 for the Constitution program. The profiles obtained in this way will hereafter be called "item profiles."

The second method by which response profiles were tabulated was similar to the first except that the time unit was the "shot." That is to say, the beginning and end of each section for which a tabulation was made were the places at which substantial changes in the visual image occurred. For example, one section began at the point at which the cameras cut from a waist shot of the teacher to a photograph of a group of Middle East refugees. The section ended when the cameras cut from this photograph to a photograph of Arab children sitting on the ground eating. There were 60 shots in the United Nations program; 55 in the Constitution program. The profiles obtained in this way will hereafter be referred to as "shot profiles."

A third method of tabulating responses was used for the Paper and Pencil group only. The GSR graphs for subjects in this group were cued for 30 second intervals, as well as for shot and item identification. This made it possible to tabulate paper and pencil interest responses and GSR emission for each 30 second interval. There were 60 such intervals in each program. The profiles obtained in this way will hereafter be referred to as "time interval profiles."

A fourth method of tabulating reponses was by individuals,

rather than by time segments. Here, for each subject the total number of clear GSRs emitted during a program were counted as well as the total number of item information segments during which he emitted at least one GSR. These will be called "individual total GSRs" and "individual item GSRs" respectively.

The 5 per cent level of confidence was used throughout this study as the criterion of statistical significance.

Results

Reliability

The average reliability of the GSR profile for a group of 30 youngsters responding to a program, when the unit of analysis is the shot, appears to be fairly high. When product-moment correlations among group GSR shot profiles were run, these were the results:

Table I

	Constitution program	UN program
Group I X Group II	.79(s)*	.82(s)
Group II X Group III	.78(s)	.77(s)
Group I X Group III	.78(s)	.84(s)

*(s) = significance.

With 58 degrees of freedom for the Consitution correlations and 58 degrees of freedom for the UN ones, all were significant.

These high and consistent correlation coefficients would seem to provide data on more than simply the reliability of this

method of obtaining response profiles. When audience studies are done using push-button or paper and pencil methods, the question is often raised whether the act of responding -- the thought and effort which must be given to a conscious response -- in itself alters the response so that it is not a very good indication of how an audience would respond if members were not pushing buttons or making marks on paper periodically. The high correlations between the GSR only group (Group III) with the paper and pencil and push button group profiles would seem to indicate that the physical acts of responding had little effect on the group "attention profile" obtained with the GSR.

Another measure of the reliability of the GSR has to do with the reliability of scoring the graphs. A sample of 27 graphs were scored independently by two coders for individual total GSRs. The correlation between the two sets of scores was .96, indicating a high degree of coder reliability.

The above measures are for inter-group and inter-coder reliability respectively. Another aspect of this question is that of intra-subject reliability. We have ample evidence that different subjects vary widely in their GSR response to films or television programs. We wanted to know whether some subjects tended consistently to emit a lot of GSRs while viewing such programs while others consistently had few changes in skin conductivity. We suspected that this was the case. However, when we took the 45 cases for whom we had clear GSR response profiles

on both programs, we found that the correlation between the total number of GSRs which each individual had during the two programs was only .26. This was not significant. The question now is whether this low correlation was due to an interaction between program content (broadly defined) and individual predispositions, or whether it was due to non-systematic extraneous factors such as how an individual felt on a given day, etc.

For the push-button and paper and pencil responses, a split-halves method of obtaining reliability was used. For each program, a response profile by shots was obtained for each random half of the subjects responding to the program by that method. Each pair was then correlated. This is analogous to the correlations among group GSR shot profiles described above. Split halves reliability obtained in this way for the push button method was .66 for the United Nations program (based upon 13 subjects in each half with correlation across 60 shots) and .69 for the Constitution program (14 subjects in each group with correlation

acorss 55 shots.) Split halves reliability for the paper and pencil method was .30 for the United Nations program (based upon 15 subjects in each group) and .29 for the Constitution program (based upon 13 subjects per group).

Thus, intersubject reliability was highest for the GSR method, almost as high for the push-button method, and by far the lowest for the paper and pencil method. Of course, given equal numbers of subjects represented in each correlation, the correlation for the push-button method would possibly have exceeded those for the GSR method.

Association among the various response measures

The GSR showed a consistent negative relationship to both the push-button and the paper and pencil responses, no matter which method of tabulating responses was used. These results are shown in Table 2. It will be noted that some correlations are marked significant (s) when other higher correlations are not. This is due to the fact that the degrees of freedom varied. To clarify the table, degrees of freedom are noted for each correlation, in addition to significance.

Table 2

	Constitution r	sig	df	United Nations r	sig	df
GSR vs push-button:						
1. By shots: All groups GSR X push-button group	-.56	(s)	53	-.77	(s)	58
2. By items: Push-button group only	-.42	(s)	31	-.30	(not s)	23
GSR vs paper and pencil:						
1. By shots: All groups GSR X p&p group	-.32	(s)	53	-.48	(s)	58
2. By shots: p&p group GSR X p&p	-.29	(s)	53	-.48	(s)	58
3. By items: All groups GSR X p&p	.31	(not s)	31	-.54	(s)	23
4. By items: p&p group GSR X p&p	.26	(not s)	31	-.47	(s)	23
5. By time intervals: p&p group GSR X p&p	-.33	(s)	58	-.38	(s)	58
6. By individuals: p&p total X GSR total (counted by 30-second interval)	.25	(not s)	25	-.14	(not s)	24

Though these correlations are not high, the relative consistency of this negative relationship, especially its consistency even when we compare the correlations for the different programs, makes it clear that the overt and covert measures either are measuring different aspects of the situation or that the GSR is an indicator of boredom rather than interest in a segment of an educational program. Previous research indicates that the former is the more likely explanation. It is conceivable that the negative relationship is an artifact of the paper and pencil and push-button response methods. The decision to

respond negatively to an authority figure -- the teacher on the screen and, perhaps, even the experimenter -- may be preceded or accompanied by a period of heightened tension which would be reflected in the GSR measure. However, a visual examination of the graphs does not reveal the clear time-order relationship between negative push-button responses, for example, and GSRs which one would expect if this explanation was valid. In addition, the consistently high correlations between the GSR shot profile obtained when subjects were responding by no other method and those obtained when subjects were also responding by either the paper and pencil or push-button method (Table I) is evidence that the act of responding by the paper and pencil or push-button method had little effect on the GSR. Another possible explanation is that subjects tend to equate "interesting" with "liking" and, with the paper and pencil or push button method, indicate a negative response for portions of a program which are difficult for them to comprehend, portions for which they must strain a bit if they are to grasp them. These would be the same sections which, because of the difficulty or strain, would be accompanied by increased GSRs. Additional research is needed to adequately test these various explanations.

As expected from the above, the paper and pencil and push button responses tended to be positively correlated. These correlations are shown in Table 3.

Table 3

	Constitution			United Nations		
Paper & pencil X push-button:	r	sig	df	r	sig	df
1. By shots	.43	(s)	53	.66	(s)	58
2. By items	-.14	(not s)	31	.61	(s)	23

These results are consistent with earlier research by this author in which correlations between paper and pencil response profiles and push button response profiles (when tabulated by 30 second time intervals) correlated significantly for both fifth grade children (.33) and for teachers (.66).[13]

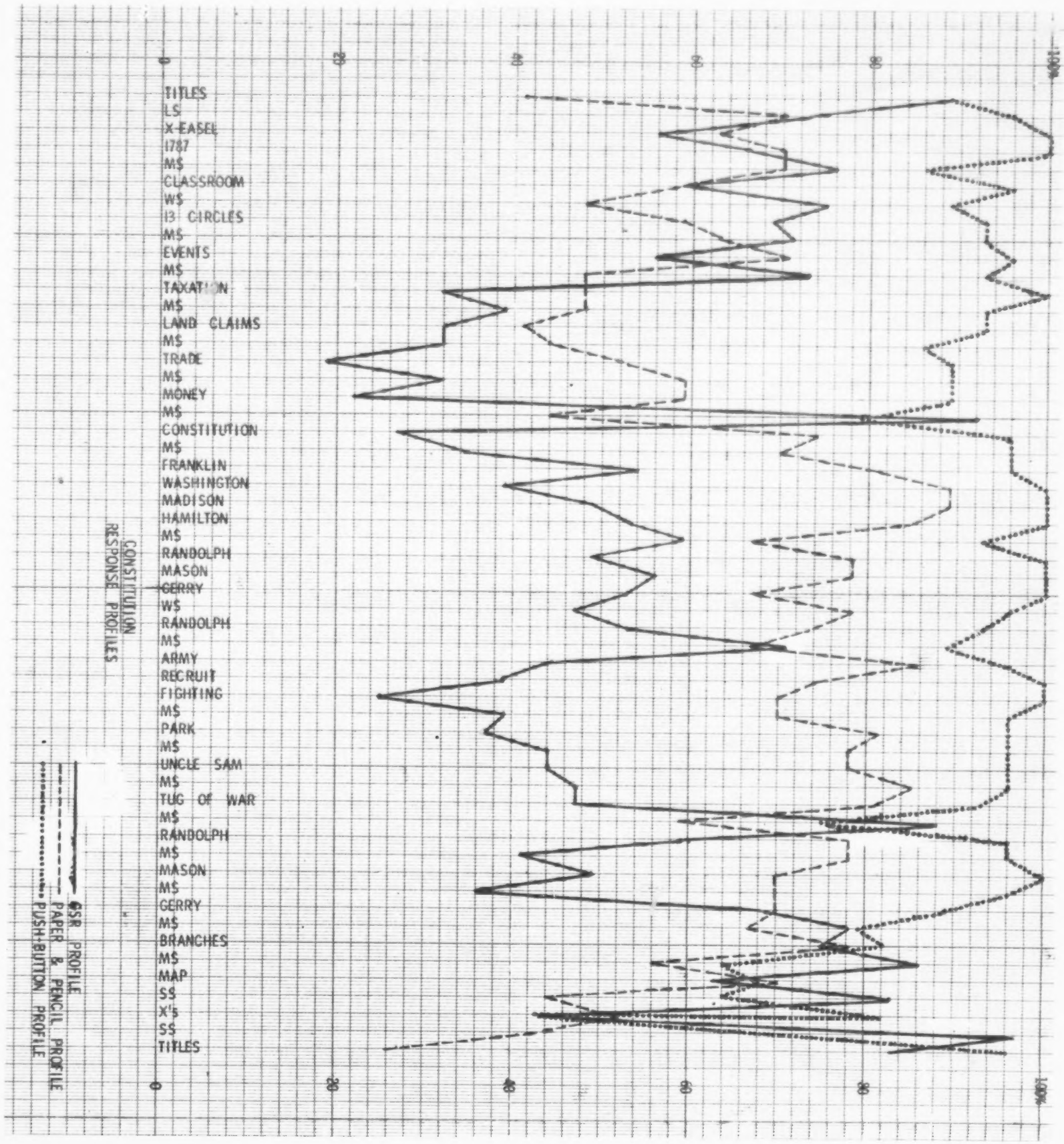
An examination of the mean for each type of response measure for each of 11 categories of shots revealed only one distinct probability of one type of shot being responded to more positively (if one can accept this directional term for the GSR) by one response method than another. The shots of the performer were accompanied by far more than the average number of GSRs per shot, but an average or less than average number of positive paper and pencil or push-button responses. This can be seen if we list the shot categories and the rank order of responses to each type of shot by each method. A 1 indicates that a shot type, on the average, was accompanied by the most GSR or the most positive push-button or paper and pencil responses.

Shot Type	Rank order of mean response rate		
	GSR	Push-button	Paper and Pencil
Charts & graphs	1	5	9
Performer (UN)	2	11	5
Performer (Const)	3	6	11
Portraits, painted	4	2	7
Photographs	5	10	4
Dates and nos.	6	7	1
Realistic drawings	7	4	6
Cartoons	8	1	8
Blackboard writing	9	9	3
Moving signs	10	8	2
Still signs	11	3	10

This higher probability of shots of the teacher being accompanied by GSRs than by paper and pencil or push-button interest responses may also be seen if one looks at a graph of these three types of response profiles for each programs. The graphs for the two programs are Figures 1 and 2 below. It is immediately apparent in these graphs that paper and pencil and push button profiles almost invariably dip for those portions of the programs in which the performer is seen. On the other hand, the GSR profile almost invariably rises.

Correlates of retention

Correlations were run between a number of measures and retention gain (difference between pre and post retention tests). None of these correlations proved significant. They are shown in Table 4.


CONSTITUTION
RESPONSE PROFILES
TITLES
LS
X-EASEL
1787
M$
CLASSROOM
W$
13 CIRCLES
M$
EVENTS
M$
TAXATION
M$
LAND CLAIMS
M$
TRADE
M$
MONEY
M$
CONSTITUTION
M$
FRANKLIN
WASHINGTON
MADISON
HAMILTON
M$
RANDOLPH
MASON
GERRY
W$
RANDOLPH
M$
ARMY
RECRUIT
FIGHTING
M$
PARK
M$
UNCLE SAM
M$
TUG OF WAR
M$
RANDOLPH
M$
MASON
M$
GERRY
M$
BRANCHES
M$
MAP
S$
X's
S$
TITLES
0
20
40
60
80
100%
GSR PROFILE
PAPER & PENCIL PROFILE
PUSH-BUTTON PROFILE

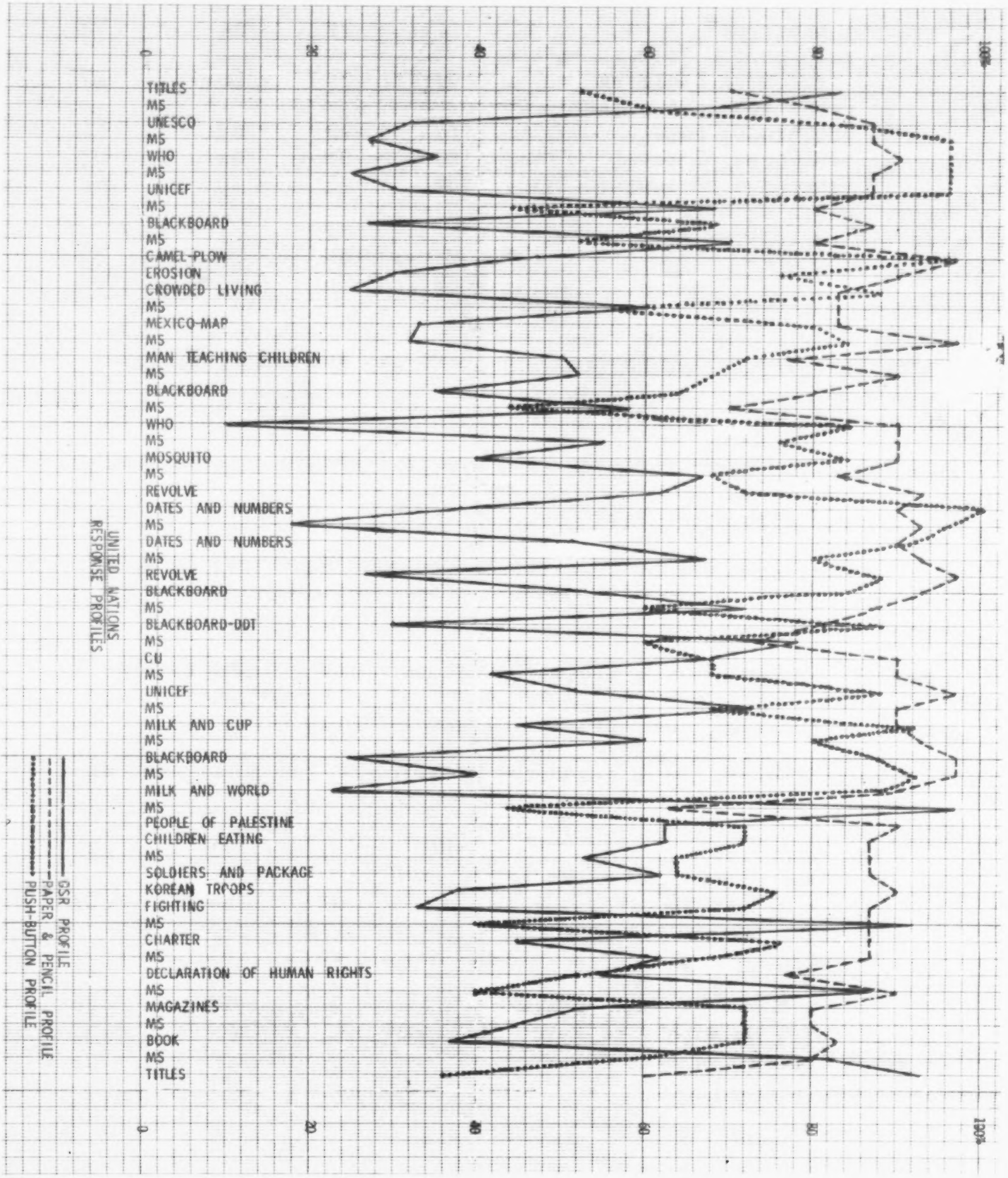

UNITED NATIONS
RESPONSE PROFILES
GSR PROFILE
PAPER & PENCIL PROFILE
PUSH-BUTTON PROFILE
TITLES
M5
UNESCO
M5
WHO
M5
UNICEF
M5
BLACKBOARD
M5
CAMEL-PLOW
EROSION
CROWDED LIVING
M5
MEXICO-MAP
M5
MAN TEACHING CHILDREN
M5
BLACKBOARD
M5
WHO
M5
MOSQUITO
M5
REVOLVE
DATES AND NUMBERS
M5
DATES AND NUMBERS
M5
REVOLVE
BLACKBOARD
M5
BLACKBOARD-DDT
M5
CU
M5
UNICEF
M5
MILK AND CUP
M5
BLACKBOARD
M5
MILK AND WORLD
M5
PEOPLE OF PALESTINE
CHILDREN EATING
M5
SOLDIERS AND PACKAGE
KOREAN TROOPS
FIGHTING
M5
CHARTER
M5
DECLARATION OF HUMAN RIGHTS
M5
MAGAZINES
M5
BOOK
M5
TITLES
0
20
40
60
80
100%

Table 4

	Constitution		United Nations	
	r	df	r	df
Variable correlated with corresponding units of retention gain:				
1. Item GSRs by individuals	.02	55	-.14	69
2. Item GSRs by items	.20	31	.00	23
3. Individual GSRs	.04	57	.01	71
4. Paper and pencil item interest profiles	.26	31	.32	23
5. Push-button item interest profiles	-.06	31	.11	23

These results again raise the question of the importance of audience "interest" when the goal is to increase the factual knowledge of the sort of captive audience which one has for in-school uses of television.

An analogous set of correlations to that described above was calculated, this time using post retention as the criterion, rather than retention gain. Results were similar, though correlations tended to be somewhat larger; enough so in some cases that they became statistically significant. These are shown in Table 5.

Table 5

	Constitution		United Nations	
	r	df	r	df
Variable correlated with corresponding units of post retention:				
1. Item GSRs by individuals	.08	55	-.19	69
2. Item GSRs by items	.23	31	.25	23
3. Individual GSRs	.12	55	-.03	71
4. Paper and pencil item interest profiles	.44(s)	31	-.09	23
5. Push-button item interest profiles	.11	31	.26	23

There are, of course, weaknesses with both sets of correlations described above. Among other things, with the latter set the correlate of prior knowledge is confounded with learning. With the former set, the ceiling effect (where many subjects already know a particular piece of information) operates to reduce the possible statistical association which may be obtained. In order to circumvent these problems somewhat, and to shed additional light on possible associations, a completely different sort of analysis was done. Taking one program at a time, we divided all the retention responses of each individual into three groups, (1) those which were correct on the pre test and correct on the post test, (2) those which were incorrect on the pre test and correct on the post test, and (3) those which were incorrect on the post test regardless of the pre test response. For each of these items for each individual then, we noted whether he emitted a GSR during the section of the program when this information was being transmitted. The same procedure was followed for paper and pencil and push button interest responses for subjects in these treatment groups. For the total group of subjects, this gave us three six-fold tables, as shown in Table 6, for each program.

Table 6
Constitution

	Correct before Correct after	Incorrect before Correct after	Correct or incorrect before Incorrect after
GSR	279 (+)	301 (+)	417 (-)
no GSR	240 (-)	251 (-)	438 (+)
p&p interest	194 (-)	181 (+)	275 (-)
no p&p interest	52 (-)	51 (-)	131 (+)
push-button interest	168 (-)	191 (+)	335 (-)
no push-button interest	26 (-)	41 (-)	90 (+)

Table 7
United Nations

	Correct before Correct after	Incorrect before Correct after	Correct or incorrect before Incorrect after
GSR	245 (+)	380 (-)	281 (+)
no GSR	212 (-)	407 (+)	243 (-)
p&p interest	127 (+)	271 (-)	227 (-)
no p&p interest	22 (-)	46 (-)	56 (+)
push-button interest	156 (+)	236 (+)	114 (-)
no push-button interest	25 (-)	46 (-)	48 (+)

Considering each six-fold table alone, a "+" sign was placed in parentheses in each cell in which the frequency was greater than one might expect by chance if learning and interest response were not related. A "-" sign was placed there if the frequency was less than expected. It is immediately apparent that there tends to be a greater probability of an interest response in the correct-correct situation or in the incorrect-correct situation and a lesser probability of this response in the correct-incorrect or incorrect-

incorrect situation. The only reversal of this was the GSR data for the United Nations program. Statistical testing of these trends is difficult because of a lack of independence among cells. The responses of a single individual may, for example, be represented in all six cells of a single table. Though the lack of independence prevents these data from meeting the assumptions of the chi-square test, such tests were made anyway in order to get some ideas of the meaningfulness of the obtained differences. In addition, it would appear that the lack of independence in this case should make it less likely, rather than more likely that significant differences would be obtained. Thus, if any are found, and the trend is consistent across programs, this would be a good indication that a meaningful relationship exists. Because there was less interest in the correct-correct category, two sets of chi-squares were run; one for each six-fold table and another for each four-fold table (eliminating the two correct-correct cells of the table in the latter case). The results of these tests are shown in Table 8.

Table 8

Chi-squares on Frequency Data in Tables 6 and 7

Response	3 x 2 table	2 x 2 table
Constitution		
GSR	5.57	4.56 (sig)
p&p	12.99 (sig)	7.65 (sig)
push button	5.49	<1
United Nations		
GSR	4.97	3.59
p&p	3.44	2.96
push button	16.37 (sig)	10.92 (sig)

For the GSRs, the single significant chi-square is probably not very meaningful because (1) there is only one out of four which was statistically significant and, (2) more importantly, the differences were in opposite directions for the two films. For the paper and pencil and push button responses, on the other hand, half of the chi-squares were significant and all differences were in the same direction for both films. In other words, in spite of non-significant product-moment correlations between interest gain and these interest measures, there may be an association between these variables. Even for the correlations (Table 4), it should be noted that the highest positive correlations were consistently between these paper and pencil and push button responses and attention gain. The consistency would seem to indicate that if more test items are used, these associations will be found to be significant.

In order to examine some of the associations between retention and other variables in still another way, the subjects at the top and bottom of the distribution of retention gain scores were isolated. In determing "cut-off" points for this purpose, an attempt was made to find natural "breaks" in the distribution. This was done independently for each of the programs. For each of the groups thus determined, mean individual GSR scores were calculated. The results are shown in Table 9.

Table 9

GSR and Retention Gain
(No. of Cases in Group is in Parentheses Following Mean.)

		High Ret. Gain	Low Ret. Gain
Mean Ind. GSRs	Const.	61.3 (6)	55.5 (6)
	U.N.	87.1 (10)	85.8 (8)

As Table 9 shows, trends for GSR were consistent across the two programs, higher individual GSR scores were consistently associated with high retention gain. None of the differences between groups on this measure was significant, however. These differences were tested for one program at a time with a t-test.

These same analyses were done using post retention, rather than retention gain, as the variable for determinging the groups. These results are shown in Table 10.

Table 10

GSR and Post Retention

		High Post Ret.	Low Post Ret.
Mean Ind. GSRs	Const.	84.6 (3)	60.7 (6)
	U.N.	81.2 (11)	58.4 (7)

We find the same association between post retention and GSR as we found for retention gain. Again, none of the differences was statistically significant.

Correlates of attitude and attitude change

There are many theories about the concomitants of attitude

change. Among these is Festinger's theory of cognitive dissonance.[14] According to this theory, attitude change tends to occur when one's prior attitudes are inconsistent with one's cognitions. Thus, when the content of a message is inconsistent with one's attitudes, one can discount or misperceive the message or one can change attitude. It would appear that the greater the dissonance, the greater the probability of attitude change. From this and similar theories (Osgood and Tannenbaum,[15] Heider[16] and others), one would expect dissonance to be anxiety-producing. Thus, a measure of anxiety should be positively correlated with attitude change.

GSR has been used by many scientists as an indicator of emotion, tension, or anxiety. It may be a useful measure for exploring dissonance theory. This study explored this possibility.

Because of the nature of the programs used in this study, it was not possible to locate particular parts of the programs which should be associated with attitude change. For this reason, the total number of clear GSR's emitted by a subject while he was watching a program was used as the measure of anxiety. When this was correlated with the amount of attitude change in the direction advocated by the program, a correlation of $.09_{(df = 55)}$ was found for the Constitution program and a correlation of $.04_{(df = 68)}$ for the United Nations program. Neither, of course, is significant. When total GSRs were correlated with the post-attitude measure

alone, somewhat higher correlations were obtained (.23 and .10 respectively) but still not significant.

As with retention gain, it was felt that these data should be examined in more than one way. Therefore, subjects were divided into three extreme groups; those who changed most in the direction of the program, those who did not change (either zero change or a change of one point), and those who changed in the direction opposite that advocated in the program. All subjects were rank-ordered according to amount of attitude change and these extreme cases were selected. Again an attempt was made to find natural "breaks" in the distributions for this purpose. Other cases were eliminated. The results are shown in Table 11. F-tests were run on the data from one program at a time to test for the signifi-

Table 11

Mean Individual GSRs by Attitude Change Groups
(o. of cases in group in parentheses following each mean.)

	High + A-Change	None	High - A-Change
Const.	101.4 (7)	82.0 (4)	43.2 (5)
U.N.	106.6 (12)	90.4 (8)	80.9 (7)

cance of differences among the Attitude Change Groups. They were not significant for either of the programs. Of course, with the small cell n's and the large variability within groups, this is not surprising. However, because of the consistency of the trend across programs, this may be worth pursuing with larger groups.

A similar analysis to that just described was done, this time determining the groups by scores on the post attitude test, rather than by attitude change. Subjects with the highest post attitude scores were compared to those with the lowest post attitude scores, again with individual GSR's as the criterion. The results are shown in Table 12.

Table 12

Mean Individual GSR's by Post Attitude Groups
(No. of cases in group in parentheses following each mean.)

	High post attitude	Low post attitude
Const.	119.3 (7)	62.0 (5)
U.N.	88.3 (8)	71.0 (7)

When t tests were run between the means for each program, neither difference was found to be statistically significant. Again, however, it is interesting to note the consistency of the trend. This is the same trend that was found when groups were determined on the basis of attitude change, though there were greater differences in mean number of GSR's between the extreme groups in that case. This similarity in results for post attitude and attitude change cannot be explained solely in terms of the same subjects being in the extreme groups in both cases. Only eight of the same subjects were represented in the extreme groups of the two analyses; that is, 26 per cent of the extreme cases in Table 9, 30 per cent of those in Table 10.

It may be hypothesized then that increased GSRs may result both from heightened interest in a message toward which you have high positive attitudes and from the anxiety-producing dissonance which precedes attitude change.

An examination of Tables 9 through 12 shows that the same trends were obtained for both the retention and the attitude measures. When this was seen, it appeared clear that we should also look at the association between retention and attitude. The results of this look are described below, in Table 13.

Table 13

Attitude-Retention Correlations by Individuals

	Constitution			*United Nations*		
	r	sig.	df	r	sig.	df
Attitude change by retention change	.24	(s)	83	.15		86
Post attitude by post retention	.39	(s)	83	.51	(s)	86

Though there is clearly a relationship between attitude and knowledge, there is obviously a good deal of variance in attitude not accounted for by variance in knowledge. This is most true and, most important for education, when we consider attitude change.

In order to assess the degree to which we were simply studying the relationship between types of individuals and these interest variables, the retention gain scores for the two programs were correlated. The same was done for the attitude change scores for the two programs. The correlation for the two sets of retention gain scores was $.44_{df = 79}$. As expected, ability to learn accounts for a large part of the variance in the retention gain scores, but not the major portion by any means. The correlation for attitude change scores over the two programs was only $-.10_{df = 77}$. Thus, we see no evidence here of a "general persuasibility" factor operating.[17]

Interest-Visual Relationships

The data from this study were examined to see whether there were any clear relationships between these measures of interest and the various types of visuals used on the programs. The results are summarized in Table 15.

Table 15
Mean Interest Responses for Differences Classes of Visuals

Shot type	n	GSR	Push button	Paper and Pencil
Moving signs	3	19.3	23.0	26.3
Still signs	7	18.9	24.9	18.4
Charts & graphs	2	42.5	24.5	18.5
Photographs	8	27.4	18.5	25.9
Realistic drawings	7	26.1	24.6	23.6
Cartoons	5	23.4	27.0	21.4
Portraits (drawings)	11	31.2	26.5	22.0
Blackboard (words & nos.)	5	20.4	19.6	26.2
Dates with nos.	2	27.0	24.0	27.0
Performer (Const.)	26	35.2	24.4	16.8
Performer (U.N.)	29	35.3	16.9	25.8

There are no clear relationships which stand out in this Table. However, shots of the performer were associated with surprisingly high GSRs. It is conceivable that the sort of tension which results in GSRs being evoked is related to seeing persons. Note that the portraits of people were also associated with a higher number of mean GSRs. This did not show up on the push-button or paper and pencil responses.

When the programs were carefully examined to see whether any particular type of visual communication or combination of visual-aural accompanied the information which was learned best or worst, again no consistent pattern could be found.

Conclusions

As shown in previous research by this author, the paper and pencil and push-button methods of audience response to films or television programs result in similar profiles. Whatever stimuli in a program evoke a positive or negative response from those responding by the push button method tend to evoke an analogous response from those responding by the paper and pencil method. When these are compared to the GSR method, however, this is not the case. As a matter of fact, correlations between GSR profiles and profiles obtained by these other methods show that GSRs tend to be evoked by those parts of the programs which paper and pencil and push-button respondents have indicated to be "boring" (Table 2). This is rather surprising, considering previous research

with GSR. One cannot onclude that GSRs are evoked more by boring than by interesting content. It seems to me that one must conclude, or at least hypothesize, that GSRs are a measure of quite a different affective response. They are evoked by a completely different dimension of content than are the more overt responses. From the present study, it would appear that these dimensions are fairly independent except when the performer is seen on the screen. These shots of the performer tend to be accompanied by an increase in tension in the audience but, at the same time, the audience rates them as uninteresting or boring.

This raises all sorts of interesting questions which can only be answered with studies in which special programs are produced in which the various content and production variables are well controlled and manipulated. Along with this, we must continue to explore ways of isolating the various types of responses. We need to be able to isolate interest from tension and, for each, be able to tell upon what part of the field the respondent was focusing -- i.e., what the specific stimulus was which was associated with the interest or tension response. We will then be able to better study and understand what sorts of stimuli evoke what sorts of responses from what sorts of audiences.

None of the response methods is a good predictor of retention from a communication. The GSR alone seems to be completely

useless for this purpose. The paper and pencil and push-button responses seem to be slightly better predictors but the correlations are still so low as to have little, if any, meaning. These results tend to substantiate those of Twyford.[18] Whether members of a captive audience think parts of a program are interesting, or whether parts of a program are accompanied by increased tension in the audience, has very little to do, apparently, with whether members of the audience retain information from the program. This, of course, has interesting implications for the producer of educational films or television programs.

GSR scores showed more relationship to attitude and attitude change than to retention. This is most clear if Tables 9 and 10 are compared to Tables 10 and 11. It seems to me that this is the most interesting and potentially fruitful finding of this study. It indicates that GSR may be a means of further exploring problems involved in persuasive communications and, most especially, the relationship of dissonance to attitude change from communication.

Footnotes

1. S. L. Becker, "Reaction Profiles: Studies of Methodology," Journal of Broadcasting, IV (1960), 253-268.

2. Ibid.

3. Ibid.

4. S. L. Becker and G. J. Wolfe, "Can Adults Predict Children's Interest in a Television Program?" The Impact of Educational Television, ed. Wilbur Schramm (Urbana: University of Illinois Press, 1960), pp. 195-213.

5. L. C. Twyford, "A Comparison of Methods for Measuring Profiles of Learning from Educational Films" (unpublished Ph.D. dissertation, The Pennsylvania State University, 1951).

6. G. L. Freeman, Physiological Psychology (New York: D. Van Nostrand, 1948), p. 118.

7. R. S. Woodworth and H. Schlosberg, Experimental Psychology (2d ed.; New York: Holt, 1954), p. 143.

8. Freeman, p. 119.

9. C. T. Mogan and E. Stellar, Physiological Psychology (New York: McGraw-Hill, 1950), p. 557.

10. R. S. Woodworth, Experimental Psychology (New York: Holt, 1938), p. 289.

11. W. E. Bingham, Jr., "A Study of the Relations Which the Galvanic Skin Response and Sensory Reference Bear to Judgments of the Meaningfulness, Significance, and Importance of 72 Words," Journal of Psychology, XVI (1943), 21-34.

12. Woodworth, loc cit.

13. Becker, pp. 266-267.

14. L. Festinger, A Theory of Cognitive Dissonance (Evanston, Ill.: Row, Peterson, 1957).

15. C. E. Osgood and P. H. Tannenbaum, "The Principle of Congruity in the Prediction of Attitude Change," Psychological Review, LXII (1955), 42-55.

16. F. Heider, "Attitudes and Cognitive Organization," Journal of Psychology, XXI (1946), 107-112.

17. I. J. Janis, et al., Personality and Persuasibility (New Haven: Yale University Press, 1959).

18. Twyford, op cit.

Appendix A

Instructions for subjects concerning measuring devices

Push button subjects

(After they complete pre-test.) All right now, while you are watching this program, we want to find out some things about your reactions to it. We want you to tell us whenever you think that the program is not very interesting -- whenever it bores you. You can tell us this by pushing this switch down whenever you are bored, and holding it down as long as you remain bored. Let it up when you become interested. Hold down when you are bored; let up when you are interested. We are also going to attach some other things to you which will tell us about whether your heart is beating faster during some parts of the program, etc. You won't have to do anything about these. Just try to forget them if possible. However, it will be necessary for you to sit as quietly as possible while you are watching the program because if you squirm around or move in any way, this messes up our measurements. Just try to relax and even when you push the switch down, try to move the rest of your body as little as possible.

(After film) Now we want you to take the two tests again which you took before the program. We want to see whether you have changed your mind about anything or whether you have learned some things which you did not know before.

Paper and pencil subjects

(After they complete pre-test.)

All right now, while you are watching this program, we want to find out some things about your reactions to it. We want you to tell us whenever you think that the program is not very interesting -- whenever it bores you. We have a special way for you to tell us this. Every thirty seconds, we will flash a number up beside the screen on which you are watching the program. This number will correspond to one of the numbers on this sheet of paper. You will see that the right side is labelled INTERESTING; the left side is labelled BORING. Check it on the right side (on the line with the right number) if you are interested; check it on the left side if you are bored. On the right if you are interested; on the left if you are bored. We are also going to attach some other things to you which will tell us about whether your heart is beating faster during some parts of the program, etc. You won't have to do anything about these. Just try to forget them if possible. However, it will be necessary for you to sit as quietly as possible while you are watching the program because if you squirm around or move in any way, this messes up our measurements. Just try to relax and even when you are marking the paper, try to move the rest of your body as little as possible.

GSR only subjects

All right now, while you are watching this program, we want to find out some things about your reactions to it, but you won't have to do a thing but relax. We are going to attach some things to you which will tell us about whether your heart is beating faster during some parts of the program, etc. You won't have to do anything about these. Just try to forget them if possible. However, it will be necessary for you to sit as quietly as possible while you are watching the program because if you squirm around or move in any way, this messes up our measurements. Just try to relax and move as little as possible.

Appendix B

Section of paper and pencil chart